TECHNIQUES FOR THE GAS CHROMATOGRAPHY – MASS SPECTROMETRY IDENTIFICATION OF ORGANIC COMPOUNDS IN EFFLUENTS

JULY 1989

Environment Ontario

Jim Bradley
Minister

ISBN 0-7729-5983-8

TECHNIQUES FOR THE GAS CHROMATOGRAPHY - MASS SPECTROMETRY IDENTIFICATION OF ORGANIC COMPOUNDS IN EFFLUENTS

Report prepared by:

R.E. Clement and V.Y. Taguchi
Laboratory Services Branch

JULY 1989 Rev.

Copyright: Queen's Printer for Ontario, 1989
This publication may be reproduced for non-commercial purposes with appropriate attribution.

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
BASIC PRINCIPLES OF GAS CHROMATOGRAPHY-MASS SPECTROMETRY	3
1. Gas Chromatography	3
2. Mass Spectrometry	7
BASIC PRINCIPLES OF ORGANIC COMPOUND IDENTIFICATION	10
1. Manual Interpretation of Mass Spectra	10
2. Computer Comparison of Mass Spectra	12
3. Use of Retention Data	13
4. Other Techniques	13
5. Selected Ion Monitoring	14
6. Quantification of Compounds by GC-MS	14
7. Factors Affecting Accuracy of Identification	16
CRITERIA FOR THE CHARACTERIZATION OF ORGANICS IN EFFLUENTS	18
1. Computer-based Library Search	19
2. Printed Reference Collections	20
3. Detailed Interpretation of Mass Spectra	21
4. Experimental Conditions for GC-MS Characterization	21
5. Reporting Characterization Data for the MISA Program	22
6. Examples of Characterization Analysis	24
REFERENCES	37

ACKNOWLEDGMENTS

The examples of compound identification were provided by Dave Wang, Steve Jenkins and Robert Kleins of the Mass Spectrometry Laboratory. Thanks are also extended to Eric Reiner,, Donna McCurvin, Helle Tosine and George Crawford for their review of the manuscript and helpful comments.

INTRODUCTION

The Municipal Industrial Strategy for Abatement (MISA) is a program of the Ontario Ministry of the Environment (MOE) designed to reduce pollution entering the environment through aqueous discharges. A unique feature of MISA is that industries are required to monitor their own discharges to ensure pollution is within regulated guidelines. This work is audited by MOE to ensure compliance.

The analysis of industrial discharges for organic contaminants present at trace concentrations is a difficult task. In some cases, hundreds of organic chemicals may be present. To detect and identify every component present in all but the most simple samples is virtually impossible and impractical. However, methods have been developed that can determine the concentrations of many compounds that are of special interest because of their potentially harmful impact on the ecosystem. Such analyses are termed "target compound" determinations because the methods used have been tailored to only those substances that have been "targeted" as being of concern. If compounds other than the specified targets are present, they may remain undetected. In the MISA program, every organic compound that is known to be present in the discharge of a particular industry will be placed on a list of target compounds that must be monitored by all industries in that sector.

In addition to monitoring for an extensive list of pre-selected target substances, an objective of the MISA program is to periodically examine discharges for new contaminants that may be present. Therefore, identification of as many compounds as possible in effluents using the best available technology will be performed. This process of identifying as many substances as possible in a sample is termed a "characterization" analysis. To perform this work, the technique of gas chromatography - mass spectrometry (GC-MS) is required. Because optimized analysis conditions cannot be pre-selected for substances that have not yet been identified, a characterization analysis cannot achieve the

same detection limits that are routinely achieved for target compound determinations. Also the accuracy in quantifying non-target analytes is less than for target compounds.

Definitive, standardized rules for identifying compounds in a characterization analysis have not been developed by the scientific community, although the general principles are well known. A series of discussions on this topic has appeared in the literature (1-4), and compound identification by GC-MS has been described in recent text books (5-8). However, in the absence of established guidelines for characterization analysis, the degree of certainty of identification of substances by different laboratories/analysts is unknown.

This report describes the system used by MOE to identify non-target organic compounds for the MISA program by using GC-MS data. This type of determination is called a "characterization" analysis. Application of this system enables the analyst to use a standard classification system for the reporting of characterization data of defined quality, yet allows the flexibility needed to use all available information to identify organic compounds in municipal and industrial effluents.

BASIC PRINCIPLES OF GAS CHROMATOGRAPHY-MASS SPECTROMETRY

The basic principles of gas chromatography (GC) and gas chromatography-mass spectrometry (GC-MS) are described in detail in recent references (5-12). An overview of these techniques is given here because interpretation of GC-MS results depends upon a knowledge of how these data are generated.

1. Gas Chromatography

All forms of chromatography involve the distribution, or partitioning, of a substance between a mobile phase and a stationary phase. In GC, the mobile phase is a gas and the stationary phase is an immobile, high molecular weight liquid which is deposited on or chemically bonded to the inner walls of a long capillary tubing. The term GLC (gas-liquid chromatography) is also used to refer to this separation technique. The capillary tubing through which the sample moves is called the chromatographic or GC column. Presently, most GC columns used for this work are manufactured from fused silica. They are generally 30-60 m in length and have an internal diameter of about 0.2 mm. By covering the outside surface of these capillary columns with a polymeric coating, these flexible fused silica GC columns are made more durable. The analysis of effluents for organic compounds requires extraction of the organics from the water matrix, concentration of the extract, separation of individual components of the organic extract by a·GC column and detection of the separated components as they are eluted from the GC column.

Complex mixtures of organic compounds are extracted from effluents by using high-purity organic solvents. The low-volatility organic compounds extracted from an effluent sample can be concentrated to a small volume (typically, 1.0 mL or less) by removing the extraction solvent through evaporation. This concentration step is necessary in order to obtain detection limits in the low part-per-billion (ppb: 10^{-6}g/L). Some compounds of concern may be more volatile than the extraction

solvent and would be lost in this process. Such substances are removed from the sample by directly purging the aqueous sample using an inert gas and collecting the purged volatile compounds on an adsorbent trap designed for this purpose. In either case, organic compounds from the sample are separated from the bulk aqueous matrix and concentrated for GC analysis.

The organic compounds are introduced into the GC column by injecting a few microlitres (μL) of the concentrated solvent extract into an injection port (non-volatile organics) or by heating the sorbent trap (volatile organics). An inert carrier gas (He, N_2, H_2), is used to sweep the extracted organic compounds, which are now in the vapour state, through the GC column.

Compounds that have different solubilities in the liquid phase of the GC column will take different times to traverse the length of the column. For a specific set of experimental conditions, the time it takes a compound to travel through a GC column is a physical property of that substance - called its retention time. Generally, higher molecular weight compounds will have greater retention times than lower molecular weight compounds. Also, substances that have a similar polarity to that of the liquid phase will be more soluble in the phase and will have greater retention times than substances less soluble in the liquid phase. Therefore, organic compounds in a mixture can be separated from each other by using gas chromatography, and the retention times of these compounds can be used to assist in their identification.

Some environmental samples are so complex that there are hundreds of compounds present in their concentrated organic extracts. There are currently no GC columns available that can completely separate all components of such complex mixtures from each other. However, in most cases the principal sample components can be individually detected. Figure 1 shows the GC separation of the volatile components of an aqueous sample, while

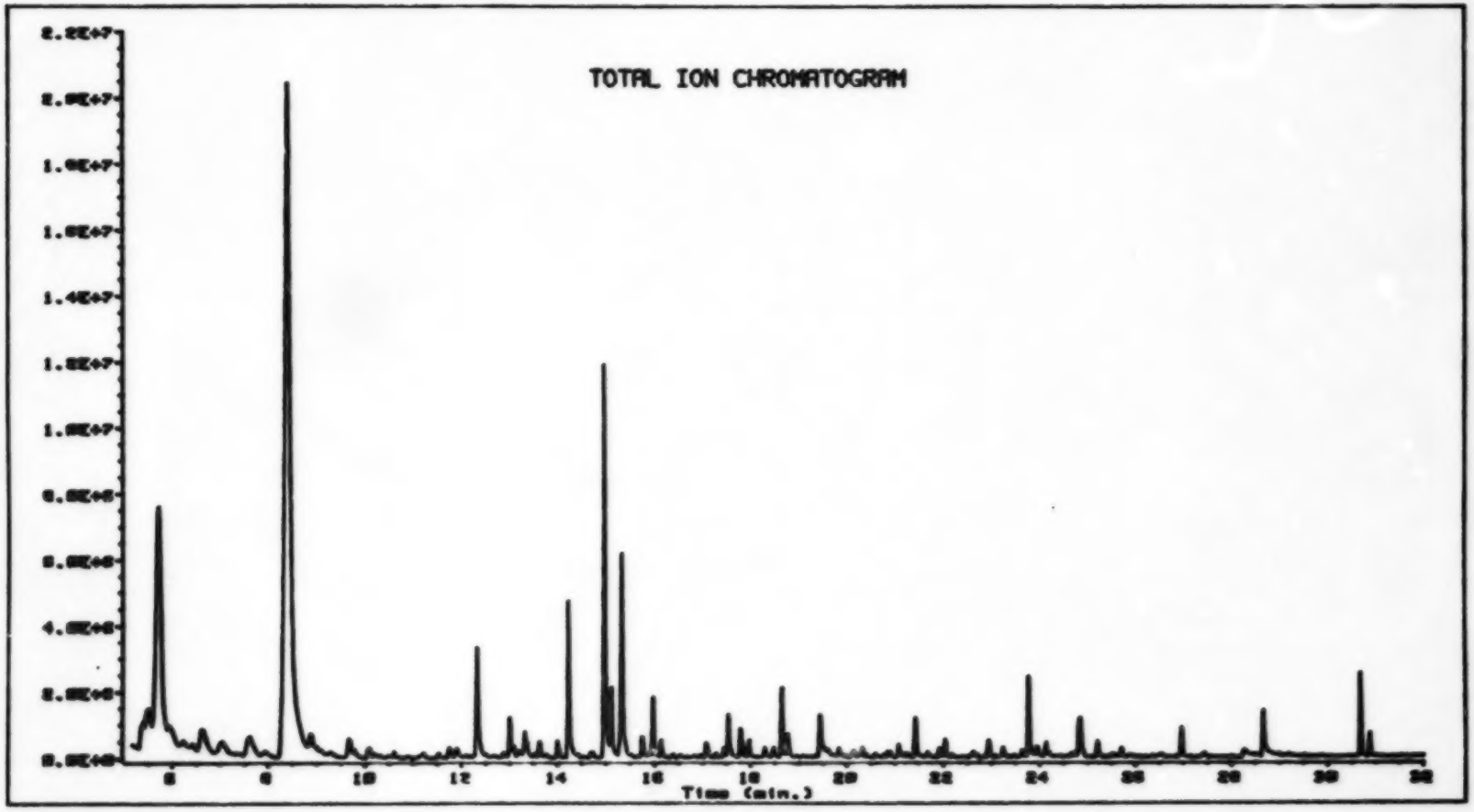

FIGURE 1. Volatile organics removed from a MISA effluent sample by purge and trap technique and separated on a 60 m fused silica capillary column.

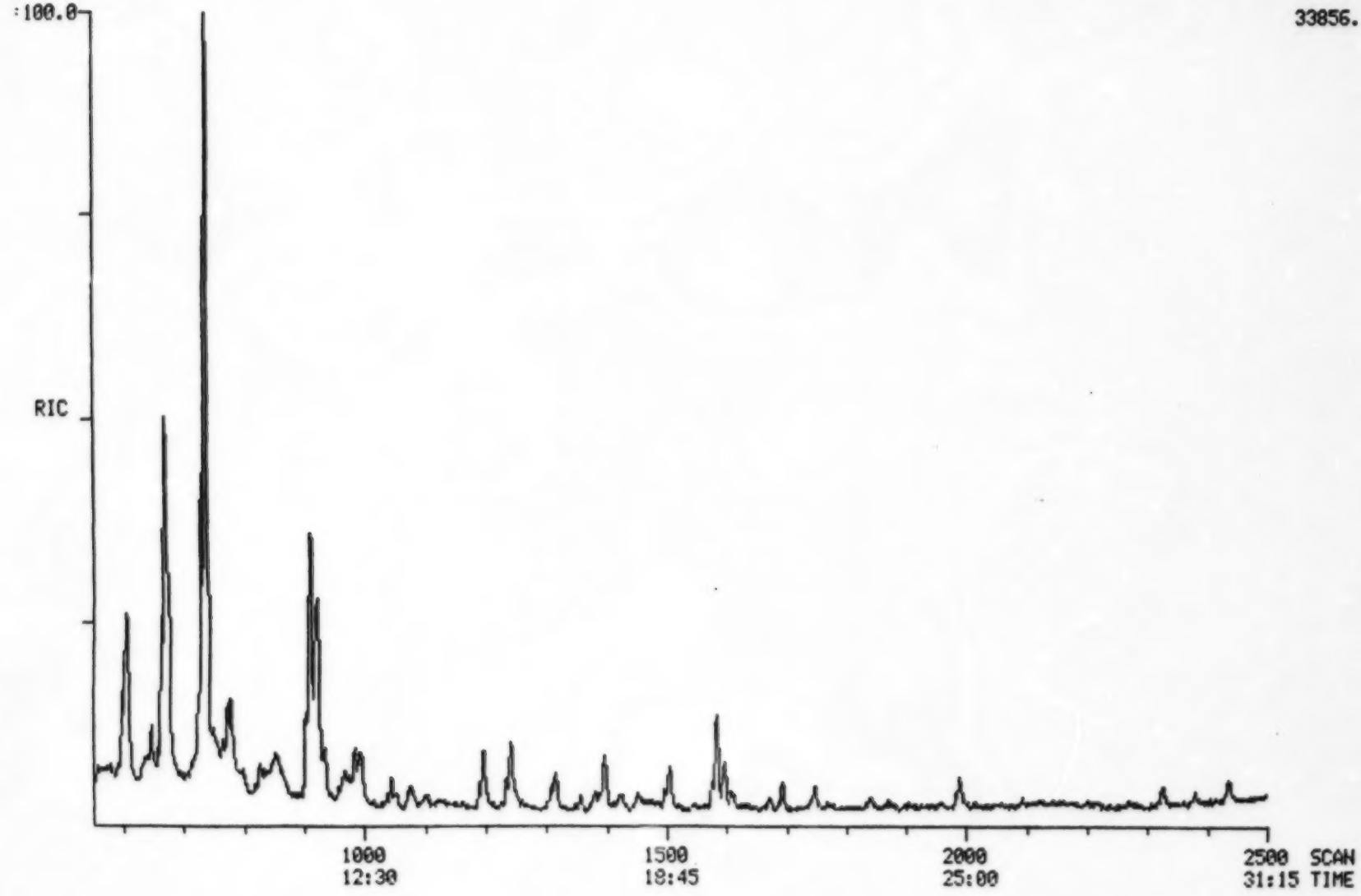

FIGURE 2. Organic components extracted from a MISA effluent by using an organic solvent and separated on 30 m fused silica capillary column.

Figure 2 shows the separation of components of a mixture of organic compounds extracted from a relatively "clean" industrial effluent using organic solvents. Although all sample components are not completely separated, the separation of the principal sample components by the GC column is good.

2. Mass Spectrometry

As the separated sample components elute from the GC column, they are monitored using any of a large number of detectors developed for this purpose. The most versatile of these detectors is the mass spectrometer. A schematic drawing of a GC-MS instrument is given in Figure 3. Initially, molecules enter the source chamber of the mass spectrometer maintained under high vacuum, where they are bombarded by electrons. The energy transferred to molecules in this process causes them to ionize and dissociate into various fragment ions. Ions may be singly or multiply-charged. The positive ions formed are made to traverse an analyzer section, maintained at 10^{-5} to 10^{-7} Torr. After the ions traverse the analyzer section where they are separated according to their mass-to-charge ratio (m/z), they are detected by an extremely sensitive device called an electron multiplier.

By plotting the abundance of ions detected versus their m/z, a mass spectrum is obtained. The mass spectrum of a compound is like a fingerprint that can be used to identify the original organic structure. It consists of a bar graph representation of the m/z of the ions and their abundances normalized to the most abundant ion (base peak). By matching the GC retention time of a sample component and its mass spectrum with those of a standard reference compound analyzed under the same conditions, a positive identification of the sample component is obtained.

Several different mass analysers have been developed. One of the most common designs consists of a square array of four parallel metal rods. By controlling radio-frequency (RF) and DC voltages to these rods an oscillating electric field is generated and this allows ions to be filtered according to their m/z. At a specific setting of voltages, only ions of the desired m/z will

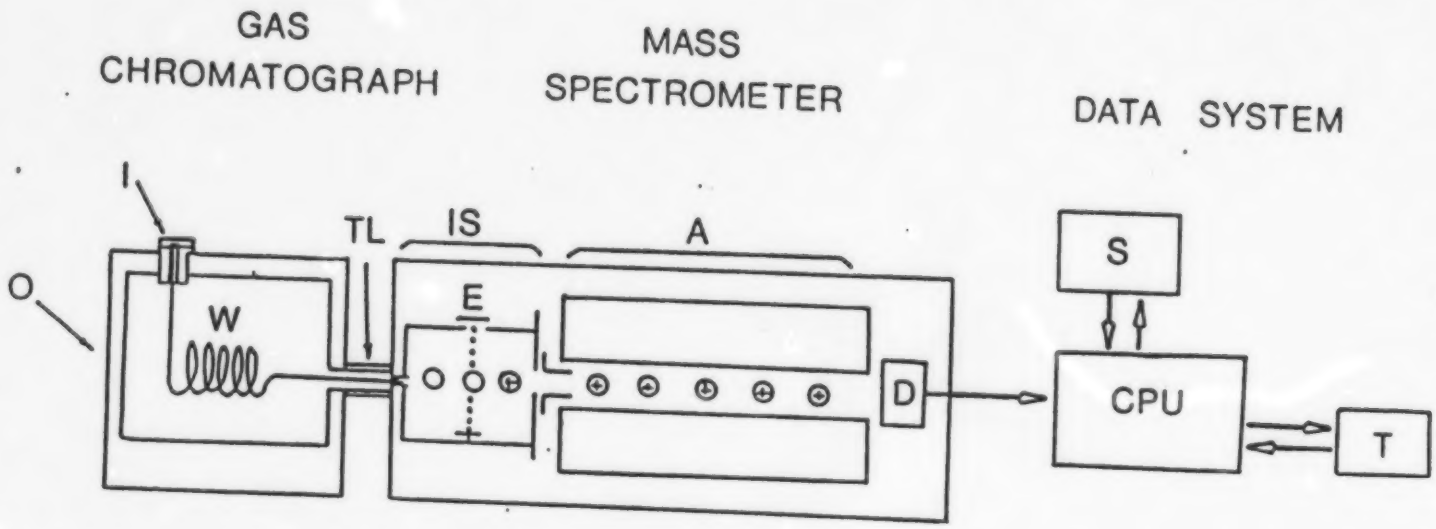

FIGURE 3. Schematic of a GC-MS System

(O=Oven, I=Injector, W=WCOT Column, TL=Transfer Line, IS=Ion Source, E=Electron Beam, A=Analyzer, D=Detector, CPU=Central Processing Unit, T=Terminal, S=Data Storage Device)

have a stable trajectory and will be able to reach the electron multiplier. By changing the applied voltages in a specified manner, the mass spectrum of a compound can be generated as the ions of various m/z are scanned. The entire process is performed in about one second. This design is called the quadrupole mass analyzer.

In another design, ions travel through a magnetic field where their motion is affected by the magnetic field strength and momentum of the ions. Conditions can be controlled to allow the analyzer to scan across a range of m/z to form a mass spectrum. Other designs are described in detail in references given previously.

An important concept in GC-MS is resolution. In GC, resolution refers to the ability of the GC column to separate components in a mixture from each other. In mass spectrometry, mass resolution refers to the ability of an analyzer to separate ions that have similar m/z. For example, quadrupole mass analyzers can resolve ions whose m/z differ by one unit (i.e. mass 16 from mass 17), whereas magnetic sector instruments can separate ions whose m/z differ by one thousandth of a mass unit or better. This is high resolution mass spectrometry. The numerical resolution for two masses that are separated is given by the formula $m/\Delta m$, where m is the nominal mass of one ion and Δm is the mass difference between that ion and the next higher mass ion that is just resolved. For example, the integer mass (or nominal mass) of both nitrogen gas (N_2) and carbon monoxide (CO) is 28. However, the actual mass of CO is 27.99492 while that of N_2 is 28.00615. A quadrupole analyzer would not be able to distinguish between these two ions (mass difference = Δm = 0.01123). By using a magnetic sector GC-MS analyzer at resolution $m/\Delta m$ = 28/0.01123 = 2,493, however, these two ions are easily resolved from each other.

Since ion m/z are not simple integers, the additional resolving power of high resolution magnetic sector analyzers is sometimes needed to improve selectivity. In addition, it is sometimes possible to establish unequivocally the molecular formula of a compound by accurate mass determinations.

BASIC PRINCIPLES OF ORGANIC COMPOUND IDENTIFICATION

A mass spectrum contains structural information of the molecule. Under a specific set of GC-MS conditions, the fragmentation pattern (mass spectrum) of a molecule depends upon specific structural features of the molecule. Therefore, it is possible to reconstruct the molecular structure from the individual ion masses observed in the mass spectrum, in much the same manner as a jigsaw puzzle is solved. It is possible for an experienced mass spectrometrist to deduce some structural characteristics of an unknown substance from its mass spectrum. However, there are so many possible organic structures that it is often not possible to assign unequivocal structures to mass spectra without comparing the mass spectrum of the unknown with a reference spectrum of the suspected structure in addition to comparing their retention times. Comparing mass spectra of unknowns with reference spectra in a data base is most effectively performed by computer search systems. Available data from other instrumental techniques and any other information (i.e., knowledge of the chemical process) should be employed where possible to supplement GC-MS data for the identification of compounds.

1. Manual Interpretation of Mass Spectra

To perform manual interpretation of mass spectra requires considerable training and experience. There is no fixed set of rules that can guarantee successful interpretations, although published books are available that describe in some detail the mass spectral characteristics of various classes of compounds (8). General steps in the interpretation of such spectra include the following:

i) Identify peaks that belong to the mass spectrum of the unknown compound. Remove other peaks which include background due to column bleed, pump oil, and other impurities.

ii) Look for the molecular ion (ion of highest mass excluding isotope peaks) - this gives the molecular weight of the unknown. This ion will not be present for all compounds.

iii) Note the general appearance - aromatic structures are more stable and give less fragmentation than aliphatic hydrocarbons which have a different, distinctive pattern.

iv) Check for characteristic isotopic patterns - if heteroatoms such as S, Si, Cl and Br are present, certain patterns of peaks in the mass spectrum will be apparent.

v) Look for characteristic neutral mass losses (i.e., loss of 35 mass units is indicative of chlorine). These neutral mass losses are usually indicative of the functional groups present in the molecule. Look at the fragment ions appearing at low m/z; these are also indicative of specific functional groups.

vi) Compare the pattern to reference spectra (i.e., homologous series of alcohols, hydrocarbons, etc. generally have similar mass spectral patterns).

vii) Interpret the spectra using the above data and any other information available.

viii) Confirm the proposed structure by obtaining the mass spectrum of the reference standard using the same GC-MS conditions.

Manual interpretation of mass spectra is tedious, requires considerable skill, and is subject to verification. Therefore, this approach should be used to suggest possible structures or to identify the compound class of an unknown only after failure of a computerized search system to provide this information.

2. Computer Comparison of Mass Spectra

Much has been written concerning the use of various computer methods for the identification of unknown mass spectra (13). Computer search systems compare mass/intensity data of the unknown spectrum with similar data in a reference file of standard mass spectra. The degree of similarity is generally expressed as some type of match factor. The unknown substance must have the same or similar structure to reference compounds that give a high match factor when their mass spectra are compared by one of a number of commercially available computer search systems.

A limitation of computer identification is that the largest reference libraries available only contain mass spectra of about 40,000 to 90,000 different compounds. This may appear to be a huge number, but is small compared to the number of substances present in the environment. Fortunately, the mass spectra of a large number of chemicals of concern in the environment (pesticides, PCBs, PAHs and many others) are well-known and are included in such databases.

Two common computer search methods are the forward search and the reverse search. In the forward search method, mass/abundance values in the unknown spectrum are matched, peak-by-peak, with those in each reference spectrum. The reverse search method matches peaks of each reference spectrum with those in the unknown. Under ideal conditions, the forward and reverse search methods are equally effective. However, the reverse search is the preferred technique when studying complex mixtures containing significant background and where individual compounds are not completely separated by the GC column. Substances that co-elute from the GC column will also have overlapping mass spectra. These compounds can be identified with the reverse search by only matching those mass/intensity values in common with each reference spectrum. Once one correct identification is made, that mass spectrum is subtracted from the unknown spectrum to give a residual spectrum that may also be successfully identified. In

practice, it is very difficult to successfully identify compounds whose mass spectra overlap significantly with other sample components.

3. Use of Retention Data

As stated previously, the GC retention time of a compound under specified conditions is constant. By matching the retention time and mass spectrum of an unknown substance with those of a standard compound analyzed using the same GC-MS conditions, positive identification is obtained. The retention time is especially useful when identifying isomers of a compound, since such compounds often produce identical mass spectra but have different GC retention times.

The retention index may also be used. Based upon retention times of standard compounds, a retention index (RI) value is calculated for each unknown. RI values are more consistent than retention times which may fluctuate depending upon small variations in experimental conditions. The use of RIs has been described in detail in other publications(14).

4. Other Techniques

Other spectroscopic techniques such as infrared (IR) and nuclear magnetic resonance (NMR) are useful for compound identification. Unfortunately, these techniques require much larger concentrations of unknowns in samples than does GC-MS and therefore are of limited value in environmental research. Recent development of fourier transform (FT) techniques and optimization of instrumental parameters has made GC-FTIR a viable technique for principal component identification in environmental samples. GC-FTIR-MS systems are now commercially available. Where possible, such techniques should be used to supplement the information supplied by GC-MS compound identification.

Specialized mass spectrometry techniques such as positive and negative chemical ionization are also available for the

identification of organic compounds (15). Advanced high resolution techniques and other methods based upon the relatively recent methods of tandem mass spectrometry (GC-MS-MS) and fourier transform mass spectrometry (FTMS) are being developed but, at this time, must be considered research areas that are not generally available for routine work. Also under development are advanced pattern recognition methods for the identification of mass spectra but, to date, these other methods cannot replace identification by conventional computerized matching of mass spectra with available reference files.

5. Selected Ion Monitoring

The technique of selected ion monitoring (SIM) is used in target compound determinations. In a SIM analysis, only a few characteristic m/z of each analyte are monitored instead of scanning the mass analyzer across the entire mass range. The advantage of this approach is an improvement in S/N for the few masses monitored and therefore detection limits can be lowered by a factor of about 100. SIM is of limited use for characterization analyses because, to achieve these lower detection limits, the characteristic masses desired must be pre-selected depending upon specific analytes of interest. When using SIM for target compound determinations, at least two or three masses characteristic of the analyte should be chosen, since the possibility of false identification of sample components is increased when the complete mass spectral fingerprint is not available. The use of SIM for target compound determinations is discussed in detail in a recent publication (5).

6. Quantification of Compounds by GC-MS

Target compounds at trace concentrations in complex environmental samples are best quantified using the SIM technique described above. For characterization analyses, quantification of the substances is of secondary importance to their identification. However, it is usually important to indicate at least approximate

concentrations of compounds identified in a characterization analysis so that some indication of the environmental significance of these substances can be made.

It is generally not possible to achieve the same quantitative accuracy during GC-MS characterization work as can be obtained by target compound determinations. The very nature of characterization work is that the compounds present in a sample are not known. Therefore it is not possible to perform the detailed calibration work that is needed for accurate quantification of specific substances until after they have been identified. In many cases, the appropriate standards may not be available. Even when standards can be prepared, there may be time-related or other factors that prevent the most accurate quantification after-the-fact.

Notwithstanding the above arguments, approximate quantification of compounds identified or partially identified (i.e., compound class) in characterization analyses is possible. Such determinations are generally referred to as semi-quantitative, because their accuracy is considerably less than that achieved for target compound determinations. Quantification is performed by comparing the response of a peak with the response of either an internal or external standard of a substance that may be in a related compound class. For example, a substance in a sample extract may be identified to be a chlorinated aromatic compound. By comparing the magnitude of the detected peak with the response factor of a chlorinated aromatic organic compound such as hexachlorobenzene, an estimate of the quantity of the sample compound can be obtained even though the quantified substance has not been definitively identified. Such estimates are usually considered to be within a factor of 10 of the true value, and are reported to one significant figure only. The accuracy of such determinations can be within a factor of 2-4 or better when a sample component is quantified by using the response factor of a standard known to be very close in structure to the unknown (i.e., 1,2-dichlorobenzene can be accurately quantified by

using the response factor of 1,4-dichlorobenzene). When the component and the internal or external standard are not in the same or related compound class, an estimate of its concentration may differ by more than a factor of 10 from its true value. Many factors affect the accuracy of such determinations, including:

i) internal or external standard calibration

ii) compound class of the unknown and of the standard

iii) GC-MS operating conditions

iv) sample complexity (i.e., degree to which unknown component can be separated from other sample components)

v) recovery efficiency of sample components from the bulk matrix.

7. Factors Affecting Accuracy of Identification

The scientific literature contains the data from a large number of studies in which characterization analyses of environmental samples were performed. In the majority of cases, identified compounds are listed in tables with no discussion of the accuracy of identifications. In fact, it may not be possible to calculate any numerical probability to such identifications. Many factors affect the accuracy of compound identifications, including the following:

i) sample complexity

ii) purity of unknown mass spectrum (i.e. no extraneous mass peaks present from other co-eluting sample components or background).

iii) concentration of unknown and GC-MS response factor (a larger than trace quantity of the unknown will give a better mass spectrum than a trace quantity - at very high concentrations, the mass spectrum may be distorted).

iv) availability of reference spectrum of the tentatively identified compound.

v) compound class/special structural features (some compounds exhibit specific mass spectral features that are easily identified).

vi) training and skill of the analyst.

vii) availability of other information such as GC retention time/index

viii) knowledge of the samples being investigated (i.e., specific information regarding industrial processes).

Although calculated numerical estimates of the accuracy of compound identifications in characterization analyses are not available, it is appropriate to express the confidence of identification in terms of how the identification was performed. For the characterization of organic compounds in MISA samples, identifications are expressed as Confirmed, Provisional, Compound Class, and Unknown, depending upon what specific information was available to the analyst.

CRITERIA FOR THE GC-MS IDENTIFICATION OF ORGANIC COMPOUNDS IN EFFLUENTS

In the current state-of-the-art, the technique of GC-MS is the method of choice for performing characterization analyses of trace quantities of organic compounds extracted from effluent samples. Positive ion electron ionization (EI) is used because of the large number of reference mass spectra available in commercially available data bases. Based upon the degree of confidence of identification, mass spectra are given one of the following labels: Confirmed, Provisional, Compound Class, and Unknown. An explanation of these terms is as follows:

1. **Confirmed identification** - an identical match is observed between the mass spectrum and retention time of the Unknown analyte compared to those of an authentic standard of the identified compound, within experimental error of the measurement system.

2. **Provisional identification**- a high level of confidence is placed upon this identification based on available GC-MS data. However, an authentic standard has not been analyzed using the same conditions as the unknown. This category includes compounds identified solely by matching unknown and reference mass spectra.

3. **Compound Class**- Sufficient information does not exist for a compound identification. However, the mass spectral data are indicative of a specific molecular structure or compound class for the analyte. This category includes the case where no reference mass spectrum is found to closely match the unknown, and sufficient additional information does not exist to make a Provisional identification. Compound Class identifications may be achieved by manual interpretation of the mass spectrum.

4. **Unknown-** Mass spectra cannot be identified without additional information. No reference spectra are found to closely match the unknown mass spectrum, and no specific structural features of a compound class are evident.

In a characterization analysis of a concentrated extract of an effluent sample, each component will be classified by the above categories based upon the available GC-MS data. To move an identification from one classification to the next higher one requires additional spectroscopic or other information. The additional information used for this purpose must be described.

To perform identifications of mass spectra using computer and printed reference compilations requires additional criteria to define an acceptable match. These criteria must exhibit some flexibility because of the many different reference collections of mass spectra available, and because the computer search systems used by various commercial suppliers of GC-MS instrumentation are different. To consider a compound for the "Provisional Identification" classification, the following minimum criteria must be fulfilled:

1. Computer-Based Library Search

   i. All computer search systems compute some type of match factor or other indication of closeness of it. This factor must be at least 80% of the value that would be calculated for a perfect match. For example, Finnigan GC-MS systems have a computer search program that calculates a forward or reverse fit of 1000 for a perfect match. The fit between unknown and reference spectra must be greater than or equal to 800.

   Some Hewlett-Packard systems use software that calculates a match factor of 100 for a perfect fit.

The calculated match factor must be greater than or equal to 80. These fit or match factors are not related to a statistical confidence level for identification.

ii Reference spectra that satisfy the above criterion must be inspected visually by the analyst. All major mass peaks (≥ 20% abundance relative to the base peak) and characteristic isotope clusters must correspond with those of the unknown mass spectrum. All major peaks in the reference mass spectrum must be present in the unknown mass spectrum at comparable relative abundances (difference between abundances of peaks in reference and unknown spectra should not be greater than about 20%). The highest mass peak in both the reference and unknown mass spectra, that is not attributed to background or a co-eluting substance, must be the same.

iii If available, the GC retention characteristics of the reference compound and unknown sample component must be the same.

2. Printed Reference Collections

Standard printed collections of mass spectra such as the Eight Peak Index or EPA/NIH data base may be manually searched for mass spectra that closely match the unknown spectrum. The same criteria as stated in 1-(ii) and 1-(iii) above also apply to visual comparison of spectra. When compilations such as the Eight Peak Index that contain only a portion of the complete mass spectrum are used, a very close match between all peaks in the reference spectrum and the unknown analyte is required before assigning a Provisional identification (i.e., a match factor such as that calculated by computer search systems would be >90% of the factor for a perfect match). The most

confident identifications require comparison of complete mass spectra. The experience of the analyst is very important in manual interpretation work which is more subjective than fully computerized spectra matching.

3. **Detailed Interpretation of Mass Spectra**

Mass spectra of unknown analytes may be interpreted by experienced mass spectrometrists to obtain the analyte identity. Such identifications will not be classified as "Provisional" unless confirmation is performed by comparing the analyte mass spectrum with that of the proposed compound. If an authentic standard is obtained of the proposed compound, and the mass spectra and retention times of the standard and the sample component are the same, then the identification can be classified as "Confirmed".

4. **Experimental Conditions for GC-MS Characterization**

Certain experimental conditions must be used when performing GC-MS identifications of organic compounds as follows:

i Samples must be extracted and concentrated only; no cleanup can be performed.

ii A characterization analysis must be performed on an appropriate blank sample.

iii Minimum scan ranges for GC-MS characterization are 35-300 daltons (volatile organics) and 40-510 daltons (extractable organics). Standard conditions should be used, including use of a standard compound for MS tuning such as PFTBA.

iv The data base used for computerized searching must be at least of the size and quality of the National Bureau of Standards (NBS) mass spectral data base (NIH/EPA/NBS Mass Spectral Data Base, National Bureau of Standards, Washington, D.C. 1986).

5. Reporting Characterization Data for the MISA Program

It is important to report specific experimental conditions when reporting characterization data, because of the wide range of mass spectral reference files and search systems available. Also, some interpretation of mass spectra is performed on a partially subjective basis, and therefore the mass spectra of compounds extracted from municipal and industrial effluents must be retained for additional evaluation or auditing when required. The key elements of a MISA characterization report are illustrated in the examples shown below. Additional information used to interpret compound identities should be reported in addition to the experimental details shown.

MISA Characterization Report

Sample No.: Date Sampled:
Date Analyzed:

GC-MS:
Column:
Temperature Program:
Internal Standard Name: Retention Time:
Analysis Type:

Entry No.	Ret. Time	Conc.	Identity	ID Class	Ref. Lib.	Match No.	Audit
1	2.56	A-400		C	NBS	850	A
2	3.19	A-10		P	EPA/NIH	825	+
3	5.63	A-50		CL	W	740	B
4	7.38	A-60		U	EPI	---	B
5	8.16	A-100		CL	MI	---	+

NOTES:

1. **GC-MS:** State make and model number

2. **Column:** State type, make, length, i.d., film thickness

3. **Temperature Program:** State conditions used

4. **Analysis Type:** VO/PT = Volatile Organics/Purge and Trap

   EO = Extractable Organics

5. **Entry No:** Component # in order of elution

6. **Ret. Time:** Retention time of the component in minutes

7. **Concentration:** A-400: A=approximate value (in ppb) to 1 significant figure

8. **Identity:** Compound name

9. **ID Classification:**
   - C = Confirmed
   - P = Provisional
   - CL= Compound Class
   - U = Unidentified

10. **Reference Library:**
    - NBS = National Bureau of Standards
    - EPA/NIH = Environmental Protection Agency/ National Institutes of Health
    - W = Wiley
    - EPI = Eight Peak Index
    - MI = Manual Interpretation
    - O = Other (specify in footnotes)

11. **Match Factor:** Numerical value from the search algorithm

12. **Audit:** to be used by MOE:

    + = identified by MOE and the industry
    A = identified by MOE
    B = identified by the industry

13. **Mass spectra of all unidentified components >10 ppb must be submitted with the report.**

14. **Mass Spectra of other components may be requested.**

6. Examples of Characterization Analysis

Figures 4-9 were obtained from a GC-MS analysis of volatile organics. Figure 4 shows an expanded section of the chromatogram in Figure 1.

The mass spectrum and the best library match spectrum of component #1 are shown in Figure 5. There is very little in common between the two spectra. A low match quality of 25% confirms this observation. No structural features are evident. Therefore, this is an unidentified compound.

The mass spectrum and the best library match spectrum of component #2 are shown in Figure 6. Although the sample mass spectrum contains some extraneous peaks, the major peaks (including isotope peaks) correspond and the match quality is high at 91%. Because the reference compound was available and its mass spectrum and retention time corresponded to those of the sample component, this is a confirmed identification.

Figure 7 shows the sample and library spectra obtained for component #3. The match quality is >80% and the major peaks correspond. The sample spectrum contains a few extraneous peaks between m/z 39-43. The sample spectrum does not contain m/z 29 because the scan range was 35-300. Masses below 35 are not monitored because there are always air leaks when a GC is coupled to an MS. These air leaks give rise to peaks at 28 (nitrogen) and 32 (oxygen) that will overload the detector. The library spectrum was obtained on a mass spectrometer using a direct inlet system and therefore m/z 29 could be monitored. Since no reference compound was available, this is a provisional identification.

Figures 8 and 9 represent the 2 best library matches for peak #4. This is a provisional identification of an isomer of trimethylbenzene. The high match qualities (>90%) indicate that either isomer is possible. This identification can become confirmed when the mass spectrum and retention time of a reference standard is available.

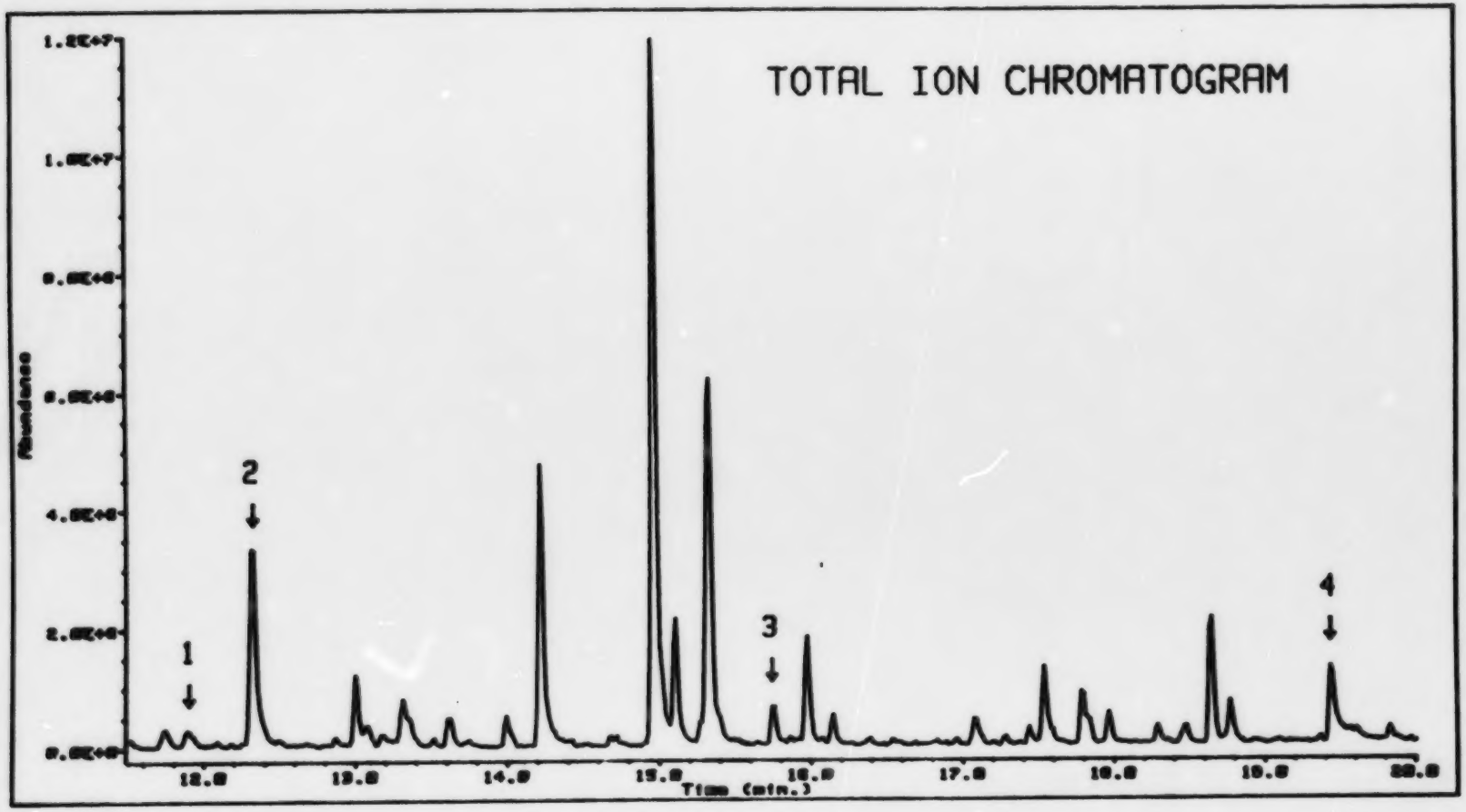

FIGURE 4. Total Ion Chromatogram of volatile organics from an effluent sample.

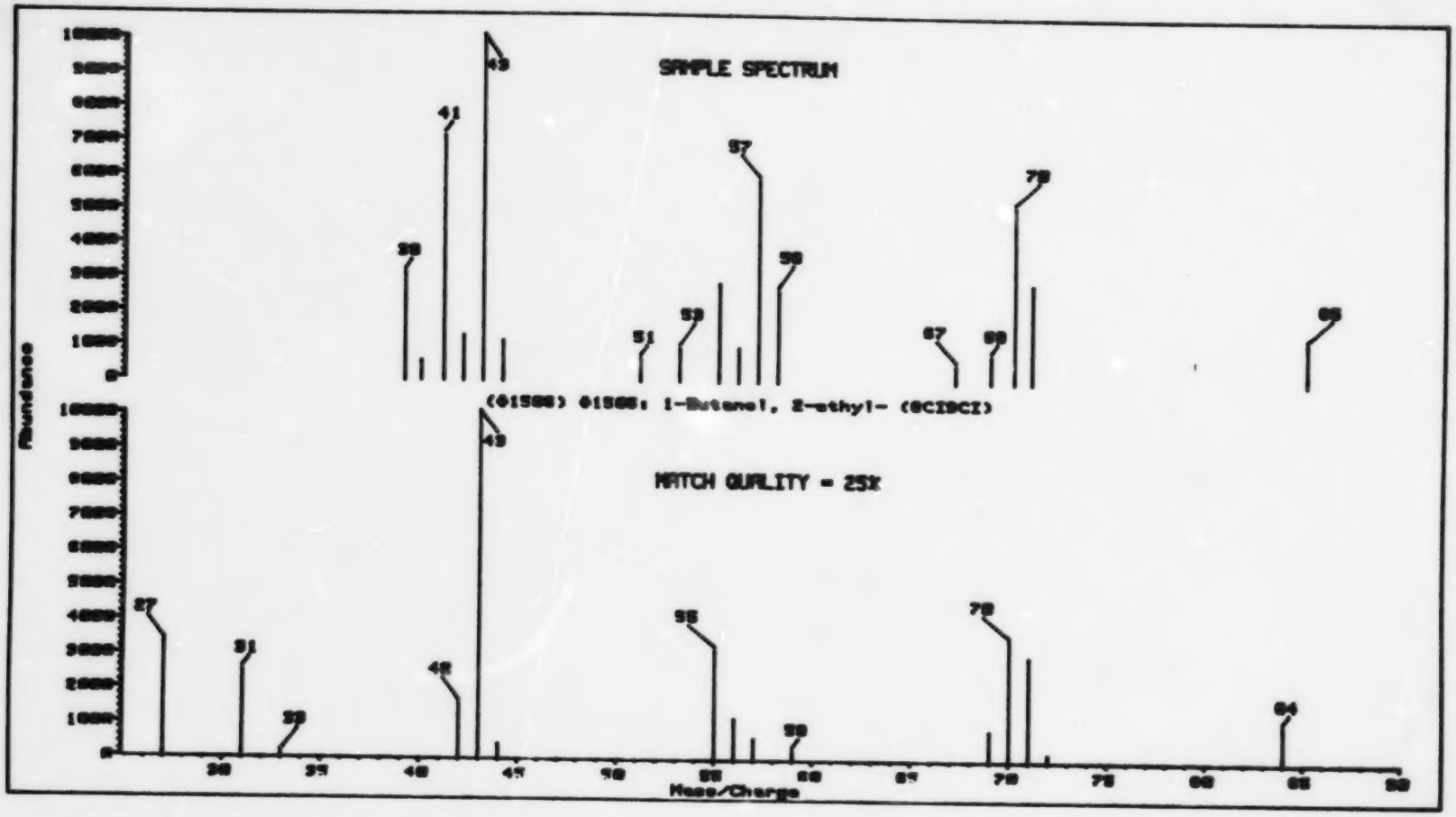

FIGURE 5. Sample and Library Spectra of Component #1.

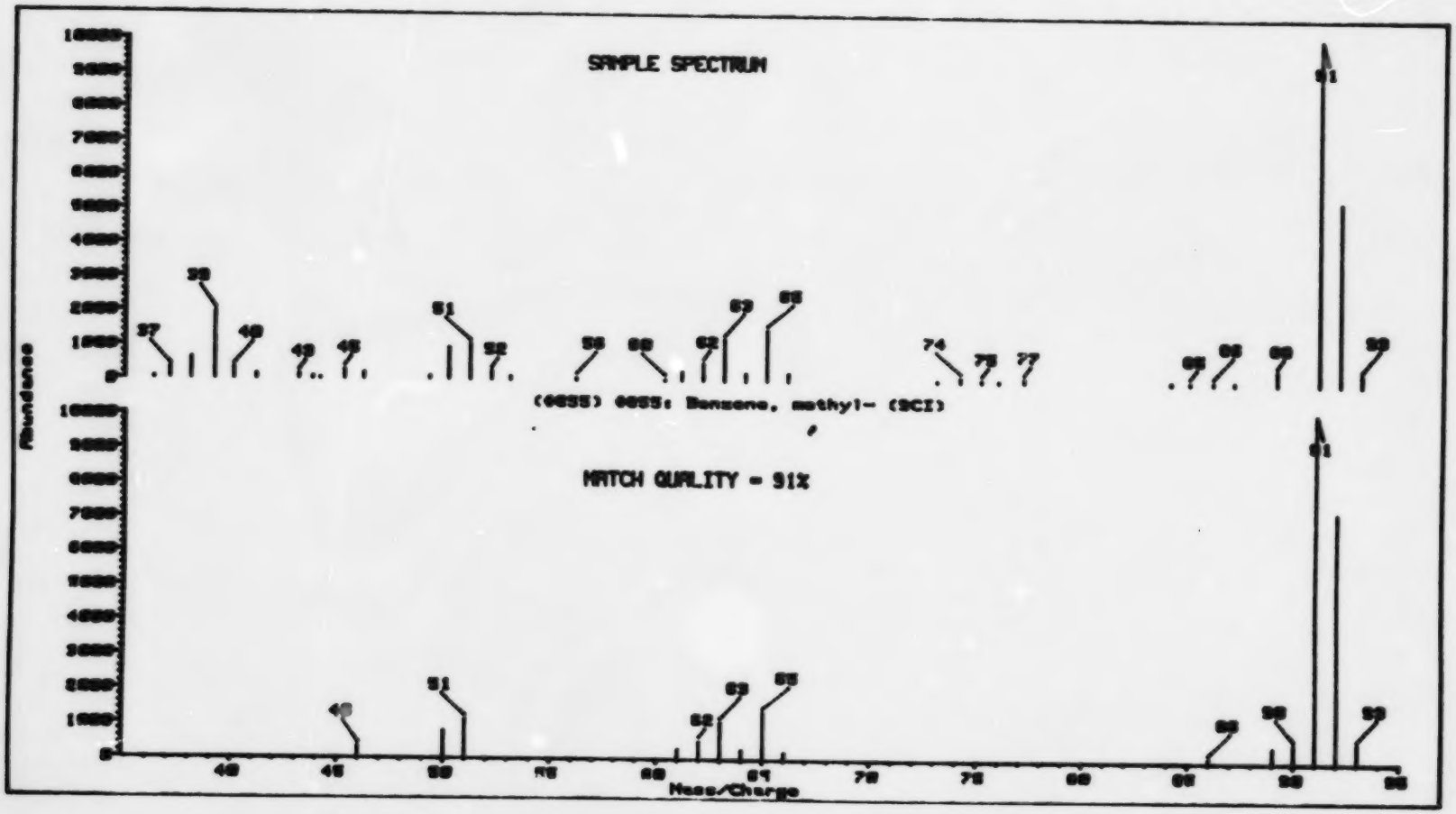

FIGURE 6. Sample and Library Spectra of Component #2.

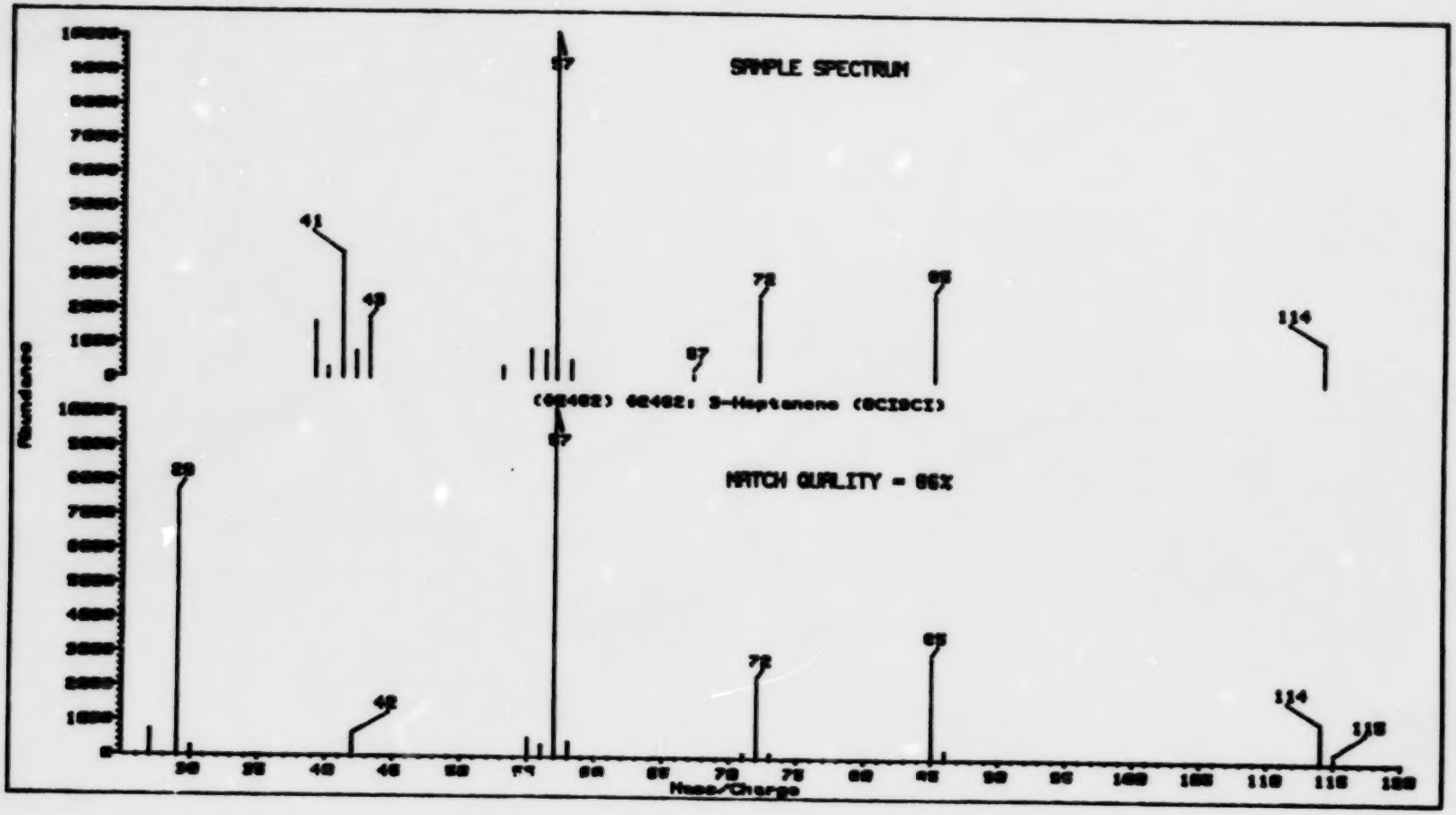

FIGURE 7. Sample and Library Spectra of Component #3.

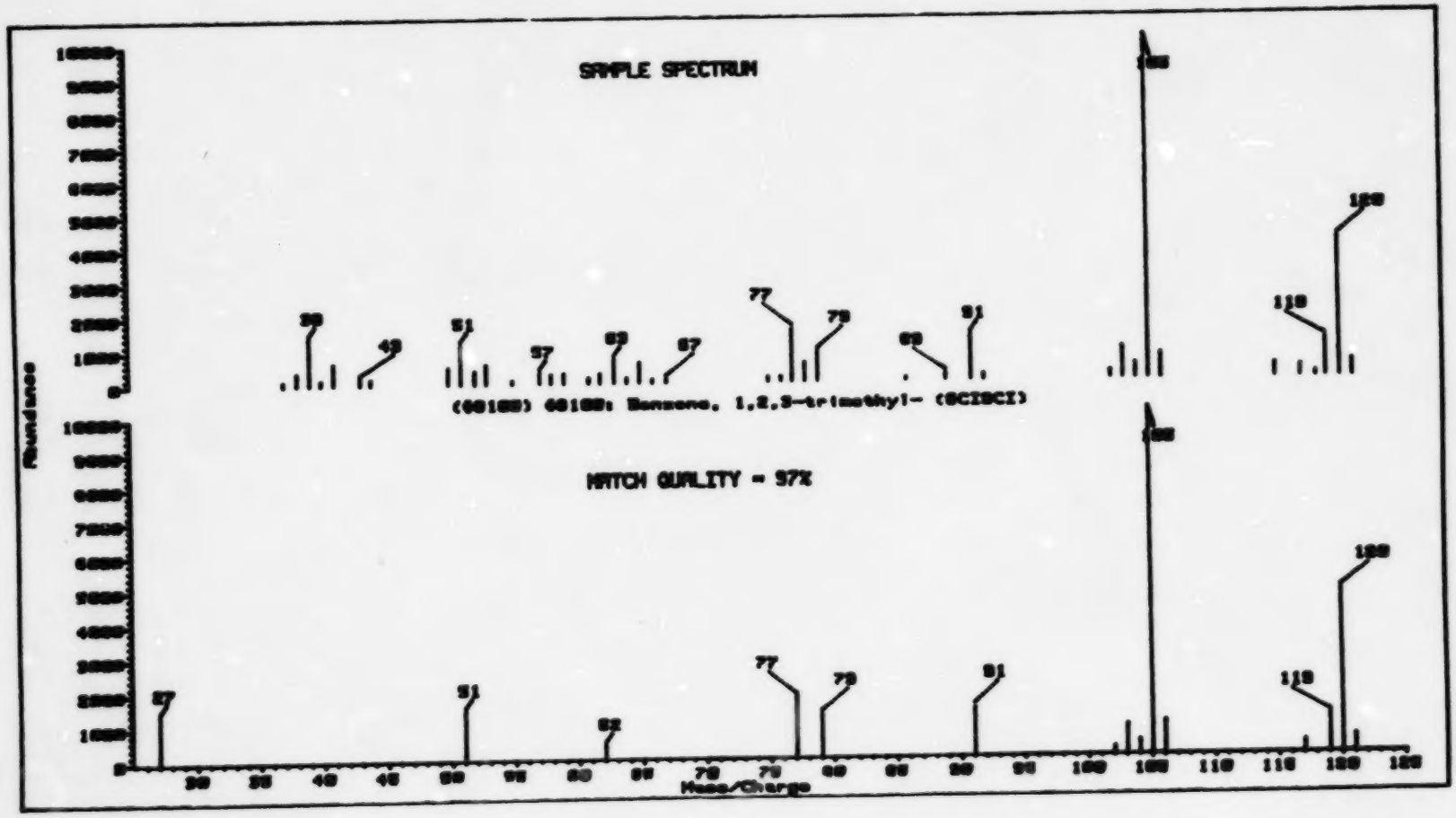

FIGURE 8. Sample and Library Spectra of Component #4.

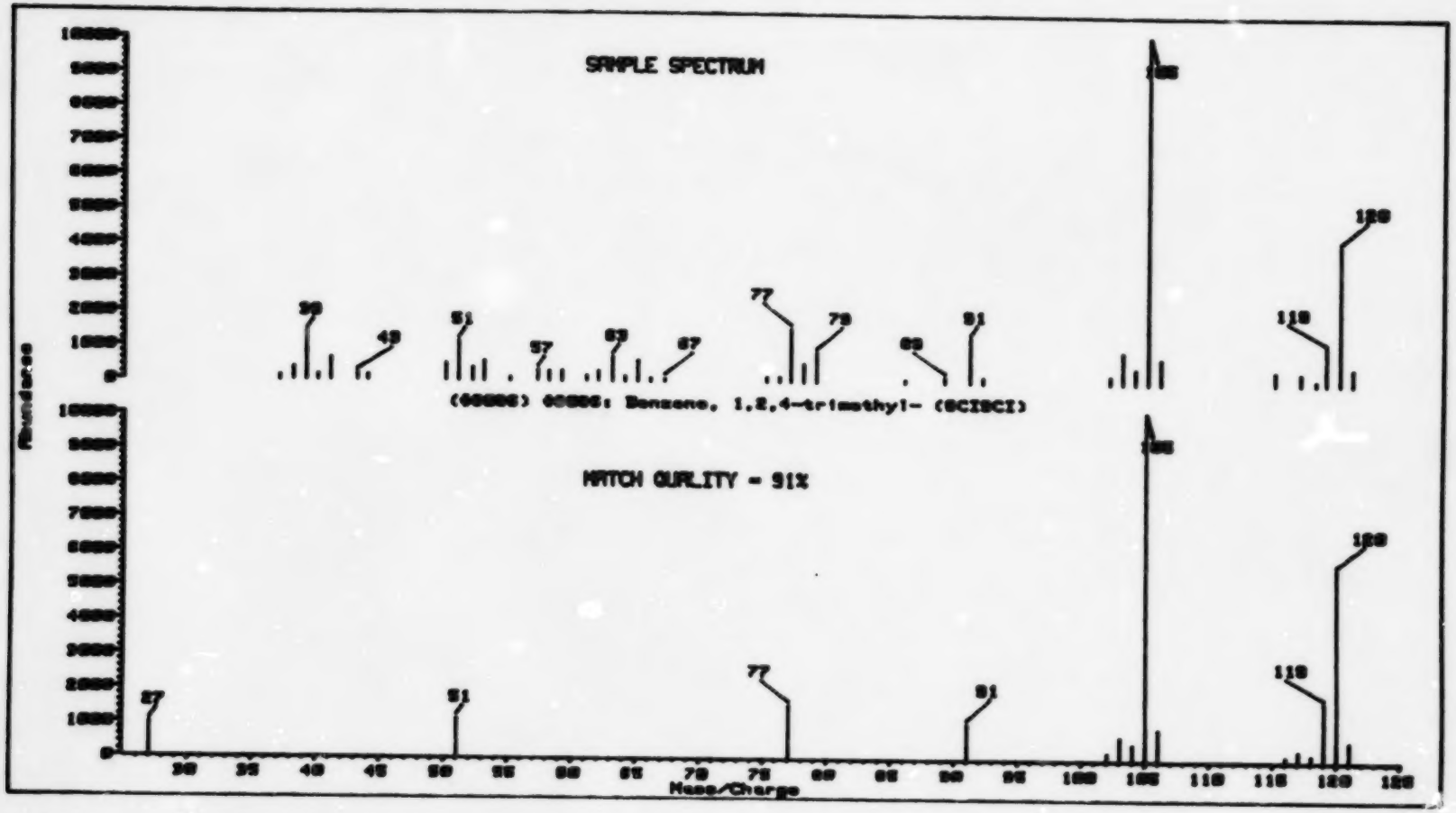

FIGURE 9. Sample and Library Spectra of Component #4.

Figures 10-14 were obtained from a GC-MS analysis of extractable organics. Figure 10 is an expanded section of the chromatogram in Figure 2.

The component at scan #268 (Figure 11) has been identified as a chloromethyl chloroethyl ether by manual interpretation (compound class). Although the best library match has a purity of 834, a forward fit of 948 and a reverse fit of 847, it does not have peaks at m/z 121 and 128. The library entry, s-dichloroethyl ether (1-chloroethyl-2'-chloroethyl ether) does not have a molecular ion at m/z 142 whereas a reference standard bis (2-chloroethyl) ether does. By analogy with these compounds, the component at scan #268 was identified as a chloromethyl chloroethyl ether. The 2 possible isomers are chloromethyl-1-chloroethyl ether and chloromethyl-2-chloroethyl ether. Since a reference spectrum was not available, this is a compound class identification.

The component at scan #606 (Figure 12) was provisionally identified as 2-oxepanone. The purity, forward fit and reverse fit were 934, 983 and 947 respectively. This identification can become confirmed when the reference standard is available for comparison of their mass spectra and retention times.

The component at scan #737 (Figure 13) was identified as a chlorine-containing compound (compound class) by manual interpretation of the sample mass spectrum. The best library match had a purity, forward fit and reverse fit of 374,813 and 407 respectively. Visual comparison of the spectra showed very little in common. A value of >800 for only 1 parameter is misleading.

The component at scan #1243 (Figure 14) was classified in the unidentified category. The purity, forward fit and reverse fit of the best library match were 555, 793 and 577 respectively. Visual comparison of the spectra showed that they were very different. No structural features were evident.

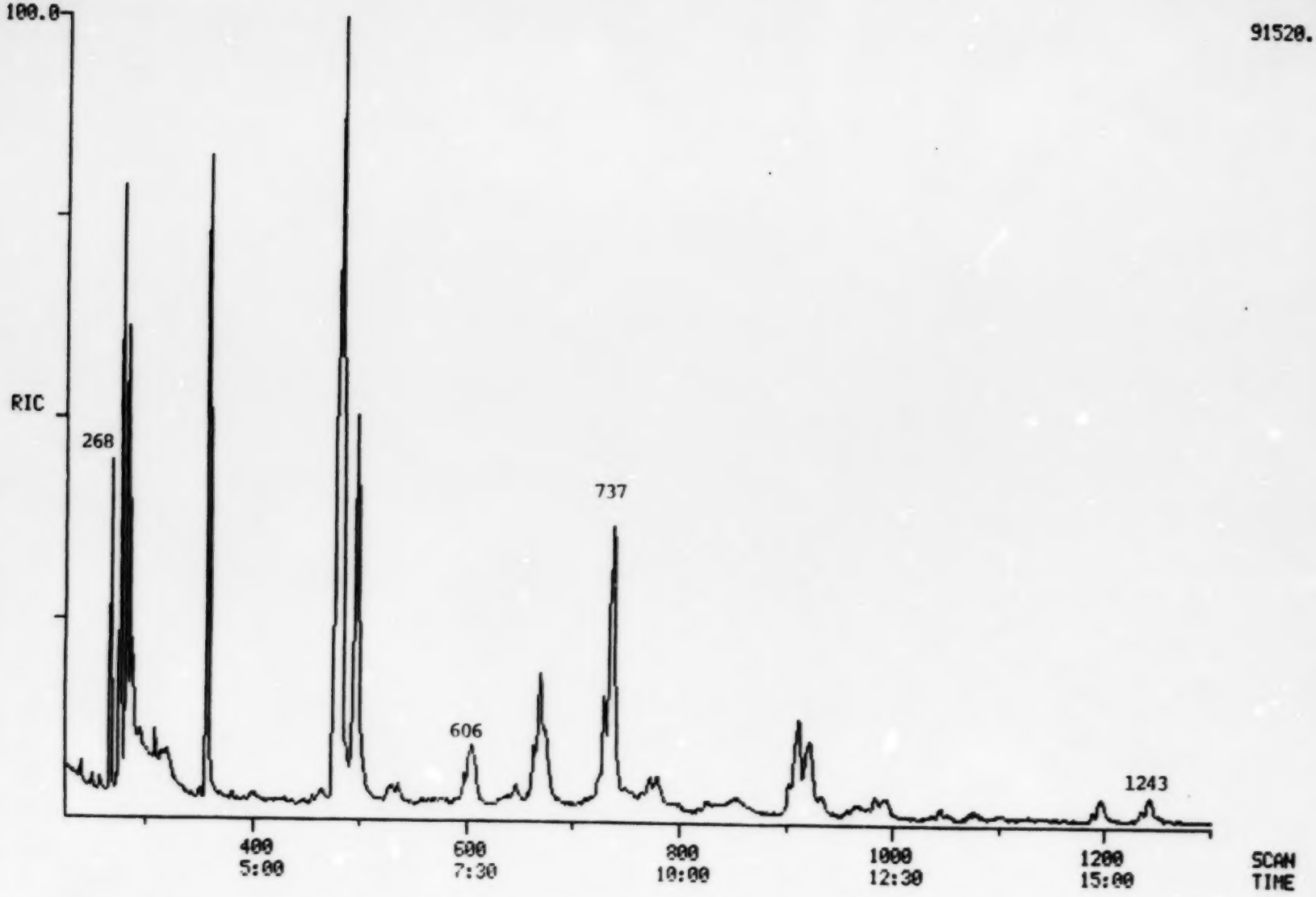

FIGURE 10. Total ion chromatogram of extractable organics from an effluent sample.

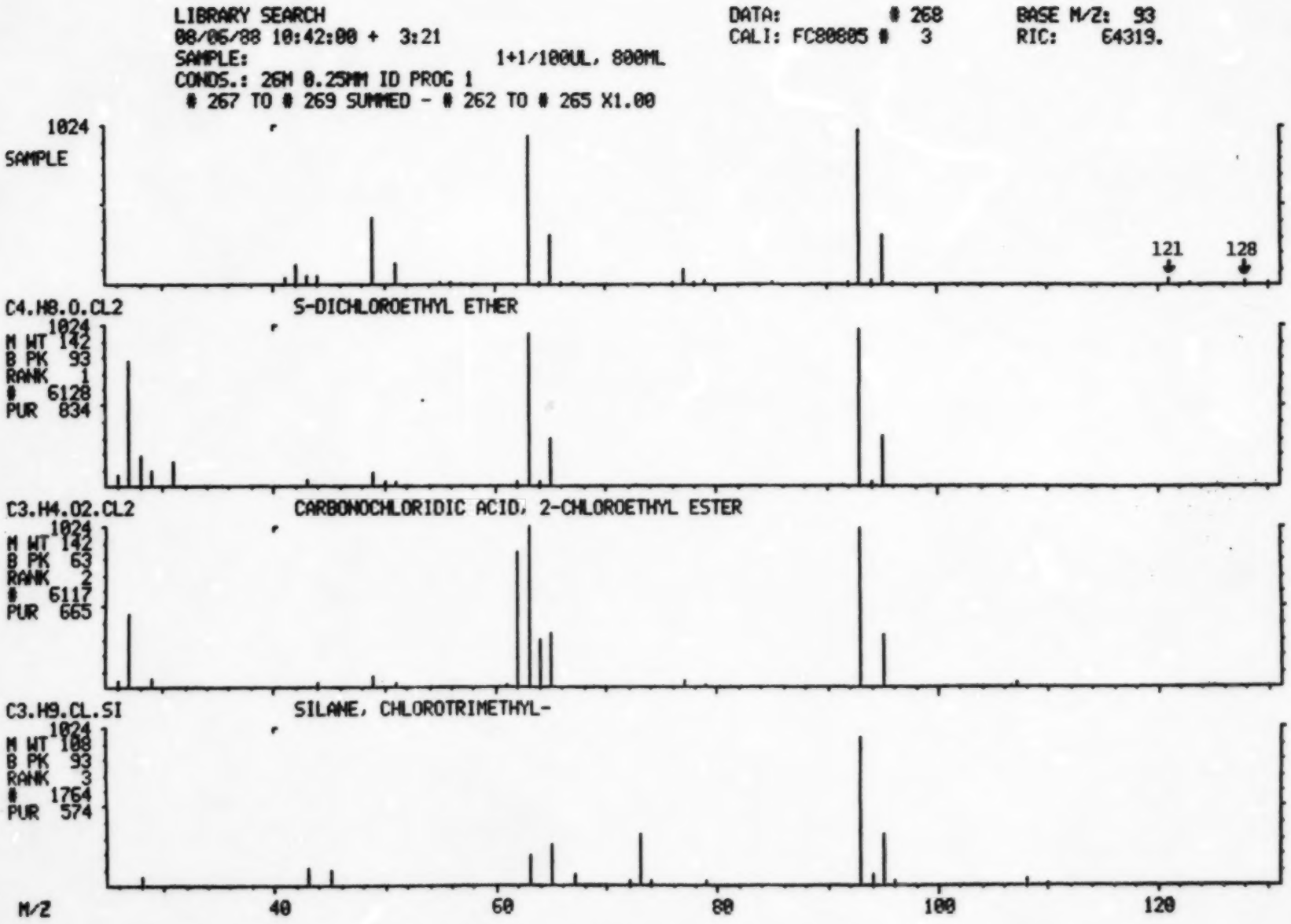

FIGURE 11. Sample and Library Spectra of Scan #268.

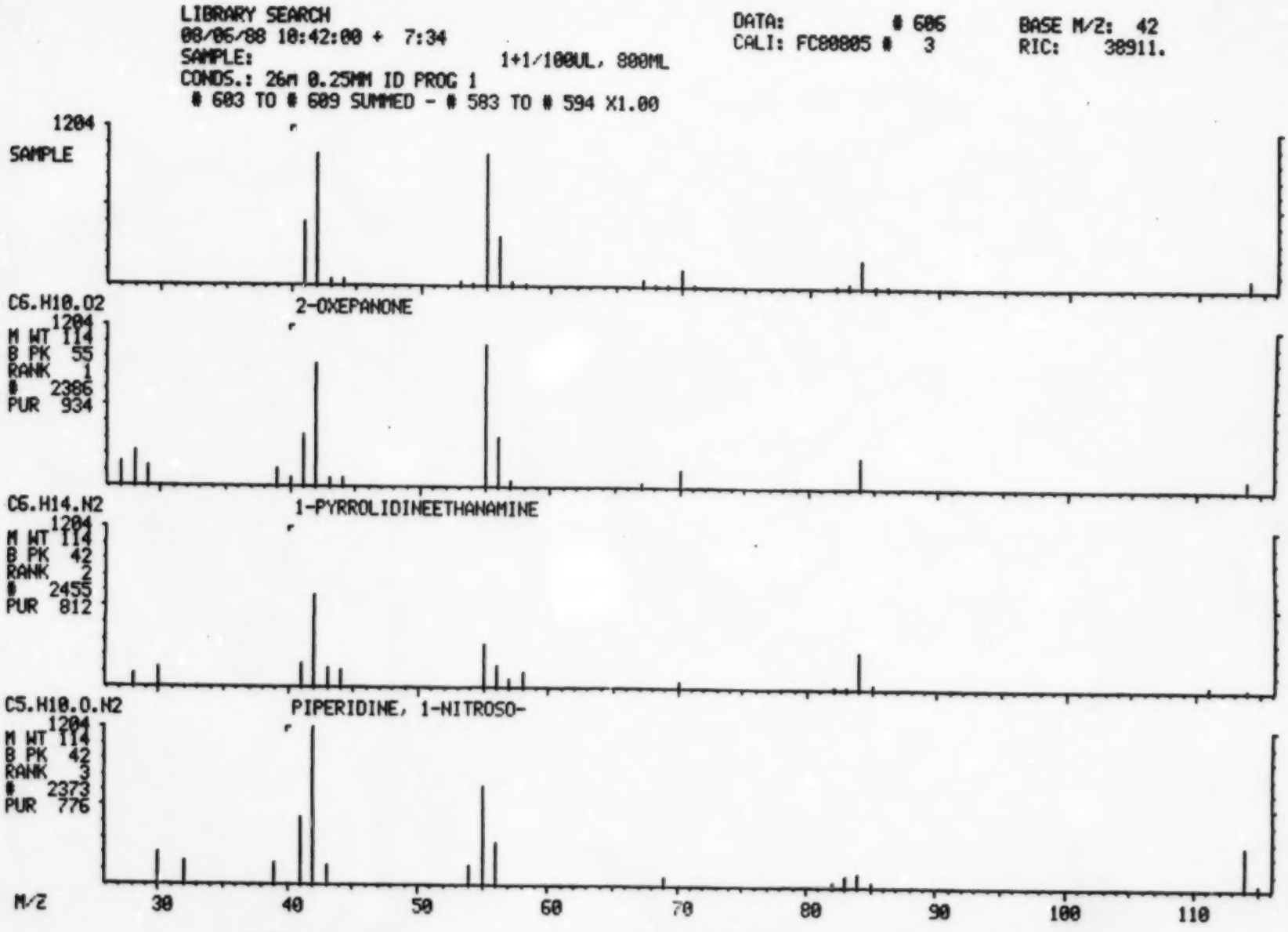

FIGURE 12. Sample and Library Spectra of Scan #606.

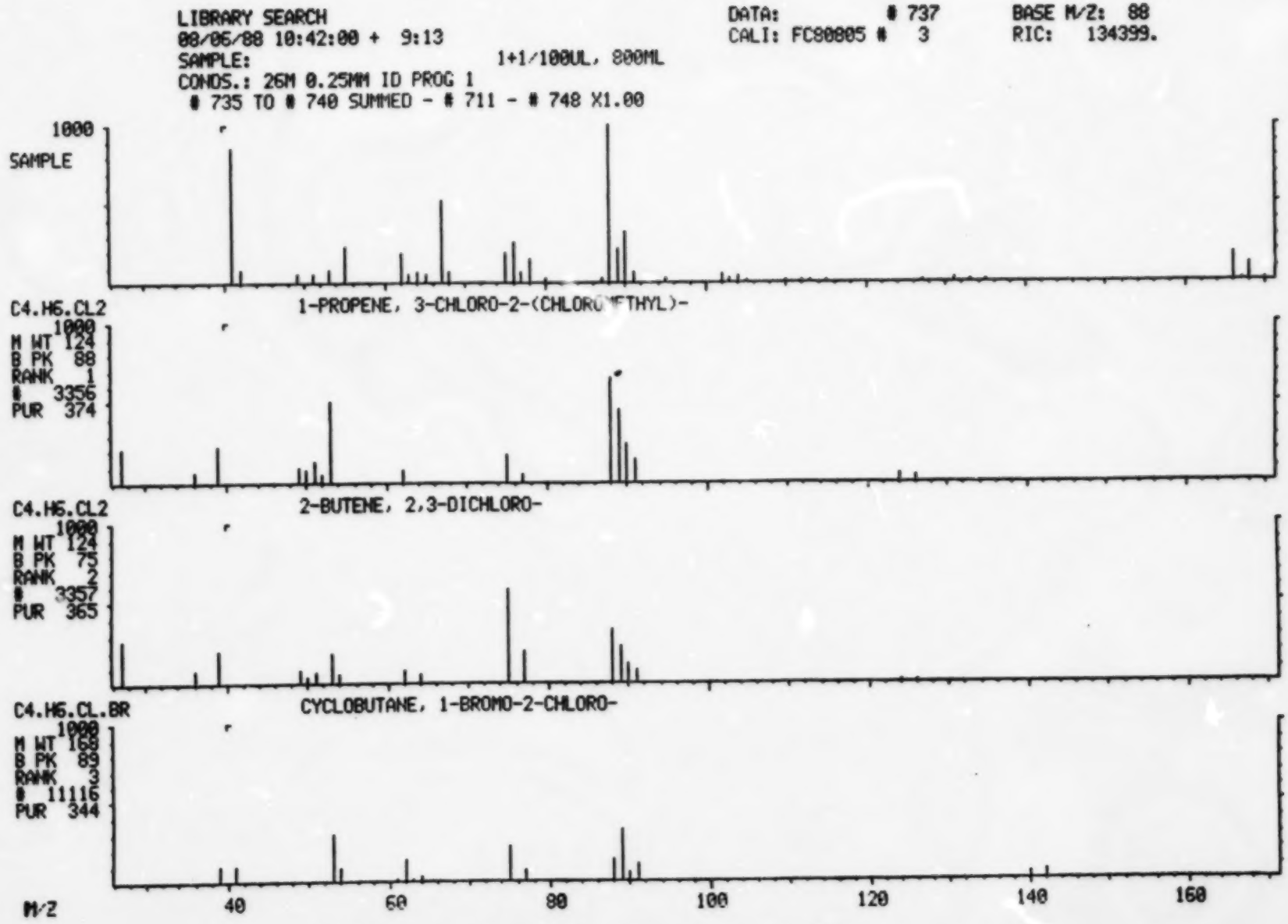

FIGURE 13. Sample and Library Spectra of Scan #737.

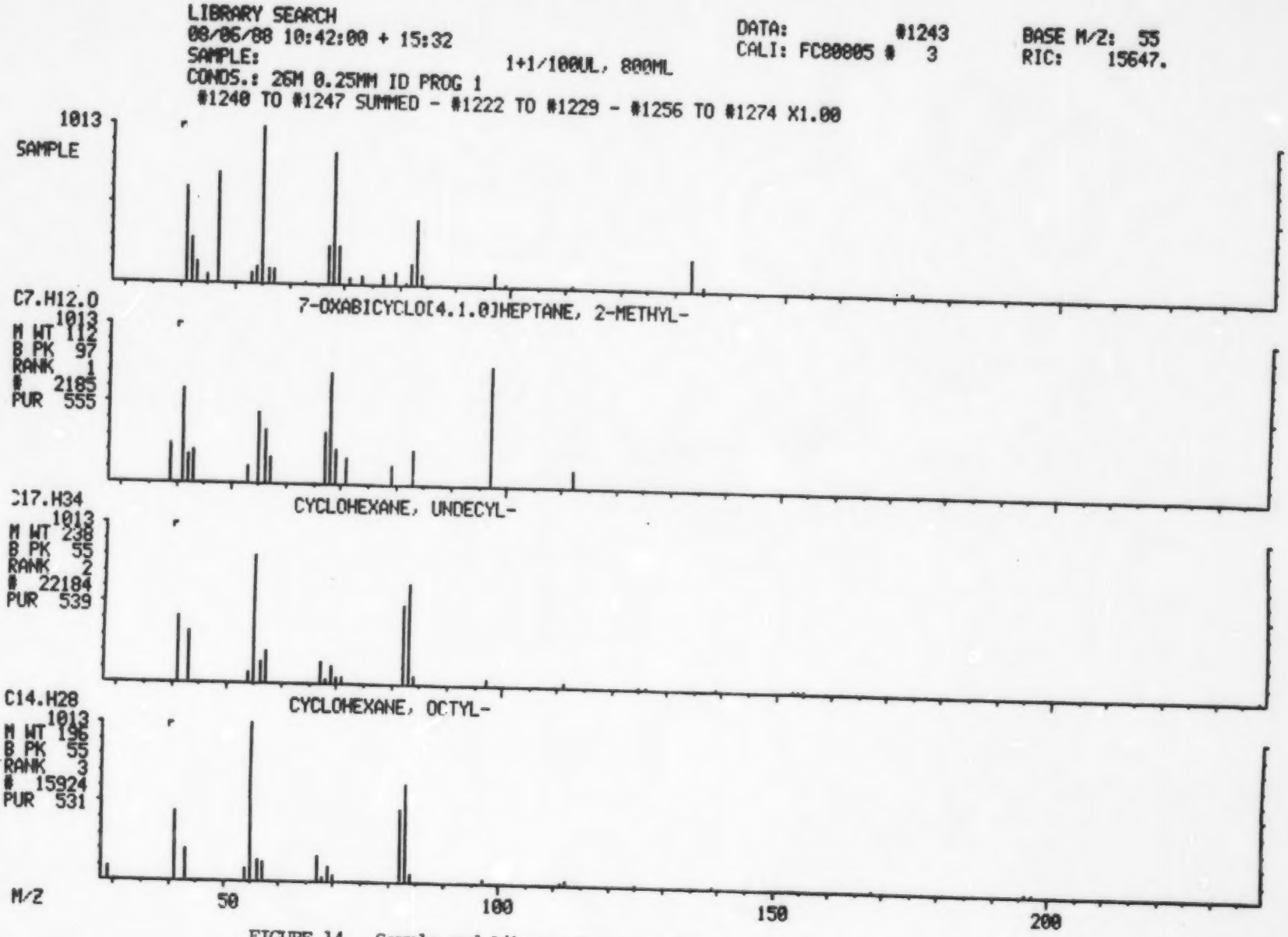

FIGURE 14. Sample and Library Spectra of Scan #1243.

REFERENCES

1. R.F. Christman, "Guidelines for GC/MS Identification", Environ. Sci. Technol. 1982, 16, 143A.

2. R.F. Christman, "GC/MS Criteria Revisited", Environ. Sci. Technol. 1982, 16, 594A.

3. Letters to the editor, Environ. Sci. Technol. 1982, 16, 595A.

4. R.F. Christman, "Editorial Policy Changes", Environ. Sci. Technol. 1984, 18, 203A.

5. Karasek, F.W. and Clement, R.E. Basic Gas Chromatography - Mass Spectrometry: Principles and Techniques, Elsevier Publishers, Amsterdam, 1988.

6. Watson, J. Throck, Introduction to Mass Spectrometry, 2nd ed., Raven Press, New York, 1985.

7. Middleditch, Brian S., Ed., Practical Mass Spectrometry: A Contemporary Introduction, Plenum Press, New York, 1979.

8. McLafferty, F.W. Interpretation of Mass Spectra, 3rd Ed., University Science Books, Mill Valley, CA, 1980.

9. Jennings, W. Analytical Gas Chromatography, Academic press, Inc. Orlando, FA, 1987.

10. Grob, Robert L., Ed., Modern Practice of Gas Chromatography, 2nd Ed., John Wiley & Sons, New York, 1905.

11. Poole, Colin F. and Schuette, Sheila A. Contemporary Practice of Chromatography, Elsevier Publishers, Amsterdam, 1984.

12. Clement, R.E. and Karasek, F.W. "Gas Chromatography/Mass Spectrometry/Computer Instrumentation", Chapter 2 in: Mass Spectrometry in Environmental Sciences, Plenum Press, New York, 1985; pp. 21-47.

13. Karasek, F.W. and Clement, R.E. "Mass Spectral Retrieval and Interpretation Systems", Chapter 7 in: Mass Spectrometry in Environmental Sciences, Plenum Press, New York, 1985; pp 123-137.

14. Budahegyi, M.V., Lombosi, E.R., Lombosi, T.S., Meszaros, S.Y., Nyiredy, S., Tarjan, G., Timar, I. and Takacs, J.M. "Twenty-Fifth Anniversary of the Retention Index System in Gas-Liquid Chromatography," J. Chromatogr. 1983, 271, 213-307.

15. Harrison, A.G. Chemical Ionization Mass Spectrometry, CRC Press, Boca Raton, FL, 1983.